SERIES ON EPIDEMIOLOGY

The Value of Randomized Clinical Trials in Ophthalmology

EMILY Y. CHEW

THE RANDOMIZED CLINICAL TRIAL IS THE GOLD standard for evaluating a therapy, providing the highest evidence for the practice of evidence-based medicine. However, not all randomized clinical trials are informative and the success of the trial in demonstrating a valid treatment or diagnostic effect will depend on the elements of the study design. The purpose of this report is to highlight the essential elements of the randomized clinical trial and how the failure to appropriately provide such elements can lead to difficulties in interpreting the data and drawing valid conclusions.

GENERAL PRINCIPLES OF RANDOMIZED CLINICAL TRIALS

THE CONTROLLED CLINICAL TRIAL IS A PROSPECTIVE study that compares the effect of an intervention, which could be a diagnostic procedure, a therapy, or a treatment strategy, with a control.[1,2] The control can be a placebo or an accepted current standard of care. It may include 2 or more intervention groups in which the treatment allocation is randomly assigned, preferably with maximum masking (blinding) for both the study subject and the treating physician. The intervention(s) may be preventive or therapeutic.

As with all clinical research, randomized clinical trials require multiple experts from different fields such as statisticians, clinical trialists, and clinicians to work together to design the best possible trial to address the research question. This research question must be timely and have sufficient equipoise that researchers are willing to accept randomization for their patients they enroll in the clinical trial. The investigator is comfortable that his or her patient may receive any of the randomized arms because there is no clear evidence for either benefit or harm. The issue of timeliness is the feasibility of studying the proposed therapy because it has not yet been widely adopted by clinicians as the standard of care.

There are a number of phases to the randomized clinical trial. Phase 1 is conducted for safety and may not be randomly assigned. It can be an escalation of drug dose to obtain the best possible dose for treatment or the maximum tolerated dose. Phase 2 studies will randomize subjects to evaluate for possible biological activity of the treatment to determine the feasibility of conducting a Phase 3 trial. The emphasis of this report is focused on the Phase 3 trials, which are designed to evaluate the effectiveness and the associated adverse effects of the new intervention. The issues to be considered include the selection of the study population (along with the sample size calculation), the allocation of the therapeutic agent or treatment strategies, the maintenance and assessment of compliance, and the outcome variables. The elimination of bias in these steps is important. The aim of this report is to emphasize the factors that are essential to a successful clinical trial design.

SELECTION OF THE STUDY POPULATION

WHEN A TRIAL IS DESIGNED, THE TARGET POPULATION often consists of the people with the highest risk of developing the disease (for prevention) or the highest risk of progressing to a more severe stage of the disease (for a therapeutic study).[3] At the conclusion of a randomized clinical trial, the investigators would like to make recommendations to the general population in whom the results of the study may be applicable. This is considered the *generalizability* of the study results. The study population of a clinical trial is clearly defined to help focus the research question and to analyze valid conclusions to such a design. In all cases, it may not include every possible scenario in the general population, thus limiting the generalizability of the recommendations for the treatment or intervention.

The study population is chosen with a careful list of predetermined inclusion and exclusion criteria to maximize the number of outcome measurements to be potentially evaluated and to standardize the screening and enrollment procedures. The feasibility of recruitment in the study population must also be considered. For example, studying a prevention therapy for age-related macular degeneration (AMD) in a population under the age of 50 may not be feasible because the development of vision-

threatening AMD usually occurs in the sixth decade and later. The cost of recruiting tens of thousands of such subjects and the duration of the study would be prohibitively expensive and daunting.

THE IMPORTANCE OF RANDOMIZATION AND MASKING

RANDOM ALLOCATION AND MASKING OF THE TREATMENT is truly the cornerstone of controlled clinical trials. When the participants are randomized, the baseline characteristics of a sufficient sample size tend to be more balanced, in both the known and unknown factors. Any observed differences between the treated and untreated groups are less likely to be attributable to chance but more likely to be attributed to the therapy. The act of randomization helps to reduce treater-selection bias and allows for statistical analyses. If there is heterogeneity in patient responses attributable to patient differences, one can consider stratifying the factors of interest. For example, in the Age-Related Eye Disease Study (AREDS), zinc was known to have little effect on cataract progression.[4] Only those participants who had either early AMD or advanced AMD in 1 eye were stratified to be randomized to zinc and the antioxidant vitamins or placebo in a factorial design, while the participants with no AMD were stratified to be randomly assigned only to antioxidant vitamins or placebo. Problems with randomization can lead to serious difficulties in interpretation of clinical trials data. For example, if randomization is conducted by having envelopes containing the random assignment at each clinical site, the envelope may be opened erroneously prior to the actual enrollment and randomization of the participant. This may result in selection bias of patients because the investigators could attempt to match the treatment assignment to patients they consider to be more suitable for the different treatment arms. This can be eliminated with centralized allocation of the treatment assignment. The generation of random numbers and the randomization scheme, and the use of "block randomization" (ensuring that there are equal numbers of the different treatment arms within a block of a certain number, unknown to the investigators), will help to ensure good randomization practice. Statistical input is crucial in designing the randomization aspect of any clinical trial.[5]

Masking (blinding) is another essential aspect of a clinical trial.[6] Ideally, participants, investigators, and the clinical trial personnel who conduct procedures for patient selection as well as the outcome measurements should be masked. This helps to eliminate the potential sources of bias in participant compliance, loss to follow-up, and investigator evaluation, improving the reliability of the results. It may not be logistically feasible to mask all the individuals involved in the studies. For example, in a study of intravitreal injections for a retinovascular disease wherein subjects are randomly assigned to an actual intravitreal injection vs a sham injection, while the patient may be masked the treating physician may not be masked. However, if the procedure is an invasive therapy such as a vitrectomy or another invasive surgical procedure, the study design is unlikely to include a sham procedure. The unmasking of the patient or the physician may have less consequence on the results if the outcome measurement is obtained by certified personnel who have no knowledge of the randomization and are able to obtain the outcome measurement objectively. An example would be the measurement of the retinal thickness on the optical coherence tomography (OCT) or progression of a disease state with color photographs, which cannot be changed subjectively by the patient.

Studies must have concurrent enrollment of the different treatment groups, including the group assigned to placebo (controls) or those in the standard-of-care group. Using historical controls would indeed invalidate the enormous benefits to a study from the process of randomization. Selection bias of participants will enter into such a situation and the historical controls may be a very different population than the one that was treated. Only with concurrent treatment groups can the comparison be valid.

MAINTENANCE AND MEASUREMENT OF STUDY COMPLIANCE

A PARTICIPANT IN A CLINICAL TRIAL IS EXPECTED TO "participate" by taking the required therapy and returning for the required follow-up visits. The success of any clinical trial depends on the success of the participants in taking their prescribed therapy and the minimization of the rates of loss to follow-up. The compliance of study participation is inversely correlated with the increasing complexity of the trial and the increasing number of treatments (number of medications, for example). The compliance issue may be tested in some study designs with a short run-in period in which the participant will be required to take a placebo for assessment of his or her potential compliance in the clinical trial. At the end of the run-in period, only those subjects who have demonstrated reasonable compliance are eligible for enrollment into the study.

OUTCOME MEASUREMENTS

THE PROCEDURE FOR MEASURING OUTCOME MEASUREMENTS needs to be standardized so that the technician who is masked is also conducting this testing with good repeatability. And if the study is conducted as a multicenter study, such standardization of procedures and certification of study personnel are crucial in reducing as much bias as possible. The outcomes need to be clinically meaningful. Surrogate measurements such as progression of diabetic

retinopathy along the Early Treatment Diabetic Retinopathy Study (ETDRS) scale rather than visual acuity loss as the primary outcomes are accepted by regulatory agencies such as the Food and Drug Administration (FDA).

SAMPLE SIZE CALCULATIONS

ONE OF THE MOST COMMON SHORTCOMINGS OF CLINICAL trials is the inadequate sample size. Power calculations are often difficult because the natural history of the disease may not be well understood and the treatment effect of an experimental intervention is often unknown. The estimates of these parameters used to make the sample size calculations are somewhat speculative at best. Rarely, the natural history data are available; the sample size of the Age-Related Eye Disease Study 2 (AREDS2) was calculated using the progression rates obtained for each of the AMD categories in the original AREDS study. The sample size calculation for the trials needs to be clearly stated in the manual of procedures and eventually reported in the manuscript. In order to support the validity of negative results from a clinical trial, a description of the sample size calculation a priori is required.

CONDUCT OF CLINICAL TRIAL

ALL RANDOMIZED CLINICAL TRIALS SHOULD BE CONDUCTED according to the standards of the International Conference on Harmonization of Good Clinical Practice (GCP). Clinical trials have been retracted in the literature when GCP is not exercised. For example, a trial in another area of medicine was considered lacking in Good Clinical Practice when auditors discovered that "there were critical deficiencies in the way patients' consent was obtained and source data were documented; the study did not have the required ethics committee (institutional review board or IRB) approval; almost all documents presented to the inspectors were copies; many existed in different unsigned and undated versions."[7] The importance of securing and properly filing the informed consent on each individual participant is emphasized.

STATISTICAL ANALYSES

THE ANALYSES OF THE DATA OBTAINED FROM A CLINICAL trial are complicated, filling up textbooks of statistics. The issues discussed here are often points of confusion in the medical community. It is essential to NOT exclude participant or observed outcomes from the analyses. Which participants should be analyzed? Most clinical trialists believe that once a participant is randomized, that participant should always be followed and included in the analyses. Exclusions refer to participants who were screened but did not qualify for the study and were not randomized. They do not bias any comparisons of intervention with the control. Withdrawals are problematic, especially if there are different rates of withdrawals for each of the treatment groups. High rates of withdrawals and imbalance in withdrawal rates would result in biased comparisons. The best solution is to minimize withdrawals from any of the treatment arms in the study.

Another point of interest in analyses is the adjustment for baseline characteristics. Because of randomization in a sufficient sample size, the baseline characteristics should be comparable in all treatment groups and the analyses may be conducted without adjustments of any variables. However, if there are imbalances at baseline values, analyses may be conducted with adjustments. This can also be accomplished if stratification by baseline variables was conducted.

THE DATA AND SAFETY MONITORING COMMITTEE

DATA AND SAFETY MONITORING COMMITTEES (DSMCS) are groups of individuals with varying expertise, including statisticians, clinicians, and clinical trialists who are totally independent of the study.[8,9] They are charged with ensuring the safety of the participants by periodic evaluation of the data. They monitor the study performance of the clinical centers, the coordinating center, or any other units related to the study. They will also evaluate baseline variables, adverse effects, and response to therapy periodically.[10] The DSMC will make recommendations to the study leadership for early termination of the study if there is unanticipated serious toxicity, greater than the expected benefits, or the likelihood of not finding a difference between the treatment and placebo (control) group.

THE IMPORTANCE OF THE RANDOMIZED CLINICAL TRIAL

THE COSTS OF CLINICAL TRIALS ARE ENORMOUS. NUMERous studies have evaluated the cost-effectiveness of clinical trials. It is important to remember that we cannot afford to NOT do clinical trials because the results may help to change how we care for our patients with diseases. It may in fact put to rest widely accepted therapies that have little or no efficacy and save both patient lives, vision, and costs to society that bear the funding of such treatments. For example, the Optic Neuritis Treatment Trial, an NIH-supported randomized clinical trial, was conducted to evaluate the role of corticosteroids in the treatment of optic neuritis.[11] Oral corticosteroids had been the unproven therapy since the 1950s. The results of the trial, conducted in the 1980s, showed that there was no role for oral steroids alone in standard doses. Instead, either no treatment or treatment with intravenous methylpred-

nisolone followed by oral prednisone should be considered for therapy. The results from the NIH-supported trials on using laser photocoagulation for treatment of diabetic retinopathy[12,13] have become the standard of care for patients with diabetic retinopathy, reducing the risk of severe vision loss by as much as 95%.[14] There was sufficient equipoise in the medical community regarding the use of laser to mount these trials. Prior to the development of laser photocoagulation, the risk of going blind after the onset of proliferative diabetic retinopathy was 50% in 5 years.[15] The results of these studies showed a remarkable and durable therapy.

The power of the clinical trial to change other medical dogmas is demonstrated in the Women's Health Initiative, a randomized controlled clinical trial on the use of hormone replacement therapy to prevent cardiovascular disease in women.[16] Observational data suggested that hormone replacement therapy reduced the risk of cardiovascular disease, but in fact, the results of the trial, which was stopped early, showed increased incident breast cancer and a trend for increased cardiovascular disease. This is an example of the issue of confounding in case-control studies, as the women taking hormone therapy were healthier and sought medical care more frequently. The use of hormone replacement therapy was erroneously interpreted as the cause of the decreased cardiovascular diseases instead of the healthy habits. The results of this trial have completely changed medical practice in women's health.[17]

Randomized clinical trials are the gold standards for determining the effects of a therapy, but in rare diseases or in studies of harmful factors such as smoking, randomized clinical trials are not feasible. Among study designs, randomized clinical trials remain unique as the only study design to evaluate the effectiveness of most interventions. Despite the associated expense, a well-conducted randomized clinical trial may have the power to change clinical practice and may, in the long run, save lives, save vision, and eventually reduce the burden of disease to both the individual and society.

REFERENCES

1. Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials Book. New York: Springer-Verlag; 1998.
2. Meinert CL: Clinical Trials: Design, Conduct and Analysis. New York: Oxford University Press; 1986.
3. Hennekens CH, Buring JE, Mayrent SL (ed). Epidemiology in Medicine. Boston/Toronto: Little, Brown and Company; 1987.
4. The Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study (AREDS): design implications. AREDS Report No. 1. Control Clin Trials 1999;20(6): 573–600.
5. Downs M, Tucker K, Christ-Schmidt H, Wittes J. Some practical problems in implementing randomization. Clin Trials 2010;7(3):235–245.
6. Psaty BM, Prentice RL. Minimizing bias in randomized trials: the importance of blinding. JAMA 2010;304(7):793–794.
7. Kleinert S, Horton R. Retraction—autologous myoblasts and fibroblasts versus collagen [corrected] for treatment of stress urinary incontinence in women: a [corrected] randomized controlled trial. Lancet 2008;372(9641):789–790. Erratum in: Lancet 2008;372(9646):1302.
8. Wilhelmsen L. Role of the data and safety monitoring committee (DSMC). Stat Med 2002;21(19):2823–2829.
9. Gallin JI, Ognibene FP. Principles and Practice of Clinical Research. 2nd ed. Burlington, MA: Epidemiology in Medicine; 2007.
10. Freidlin B, Korn EL, George SL. Data monitoring committees and interim monitoring guidelines. Control Clin Trials 1999;20(5):395–407.
11. Beck RW, Cleary PA, Anderson MM Jr, et al; Optic Neuritis Study Group. A randomized, controlled trial of corticosteroids in the treatment of acute optic neuritis. N Engl J Med 1992;326(9):581–588.
12. The Diabetic Retinopathy Study Research Group. Photocoagulation treatment of proliferative diabetic retinopathy: clinical application of Diabetic Retinopathy Study (DRS) findings. DRS Report No. 8. Ophthalmology 1981;88(7): 583–600.
13. Early Treatment Diabetic Retinopathy Study Research Group. Early photocoagulation for diabetic retinopathy. ETDRS Report No. 9. Ophthalmology 1991;98(5 Suppl): 766–785.
14. Ferris FL 3rd. How effective are treatments for diabetic retinopathy? Commentary. JAMA 1993;269(10):1290–1291.
15. Beetham WP. Visual prognosis of proliferating diabetic retinopathy. Br J Ophthalmol 1963;12(10):389–397.
16. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. JAMA 2002;288(3):321–333.
17. Enserink M. Women's health. The vanishing promises of hormone replacement. Science 2002;297(5580):325–326.